REAL TIME SUPER-RESOLUTION FOR LOW RESOLUTION IMAGES USING SRGAN

Bheshaj Kumar Dewangan

MTech AI, INDIAN INSTITUTE OF SCIENCE

ABSTRACT

GANs have been shown to obtain high quality results in image enhancement tasks; however, to obtain natural quality results high capacity large generators are usually employed, resulting in high computational costs and processing time. In 2016 SRGAN [1] was proposed for performing realistic super-resolution of images. In this project I have tried to enable real time super resolution for upsampling low resolution images using SRGAN architecture. Instead of residual blocks, inverted residual blocks are employed for parameter efficiency and fast operation. The idea is somewhat inspired by *Galteri et al* [2] for real time image enhancement. The proposed architecture named REAL-TIME-SRGAN and SRGAN were trained and compared in terms of number of parameters and other output metrics. From the results it can be inferred that real time upscaling can be done with very less parameters close to SRGAN results.

Index Terms— GANs, Super-Resolution

1. INTRODUCTION

Real time image enhancement is the task of generating high quality frames from low resolution frames. SRGAN [1] was proposed to generate super-resolution images but the architecture is not suited for real time video generation. [2] performed real time enhancement earlier in 2019 on compressed low resolution videos.

2. TECHNICAL DETAILS

2.1. Architecture Details

SRGAN [1] uses 64 filters with kernel size 9 while I have used 32 filters with kernel size 3. The residual block structure along with proposed residual block structure is shown in Table 1. I have used 16 residual blocks in SRGAN and 6 blocks in Real-Time-SRGAN. The rest of the architecture was same as that of SRGAN [1].

2.2. Learning rate Schedule

Both Generator and Discriminator had exponential decay with initial learning rate of 10^{-4} with stair case.

Residual Block Architecture	
SRGAN	**Real-Time-SRGAN**
Conv2D BatchNorm Prelu Conv2D BatchNorm	Conv2D (k=1) BatchNorm Relu DepthWiseConv2D (k=3) BatchNorm Relu Conv2D (k=1) BatchNorm

Table 1. Residual Block Architecture

Number of Parameters	
SRGAN	**Real-Time-SRGAN**
1554k	112k

Table 2. Number of Parameters

2.3. Loss function

Both SRGAN and Real-Time-SRGAN generator networks were pretrained on MSE loss. I used the same loss function (perceptual loss) as that in SRGAN i.e., weighted combination of Adversarial, MSE and VGG feature loss as given in [1].

2.4. Dataset

DIV2K bicubic dataset was used in all training experiments. It contains 800 training and 100 validation images both in low and high resolution.

2.5. Training Details

I have used upscale factor of 4 only. In both models Generator was pretrained with 500 epochs with MSE loss. For comparison randomly cropped image with size 256 is considered for training both GANs. Both models were trained for 1000 epochs.

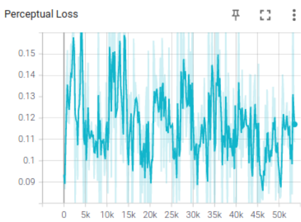
3. RESULTS

The training loss plots for Real-time-SRGAN is shown in Fig.1. The number of parameters of both models are

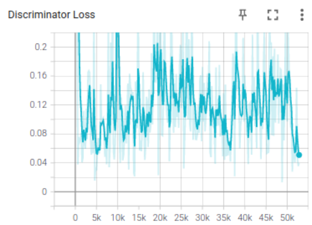
	SRGAN	Real-Time-SRGAN
SSIM	0.716	0.723
PSNR (dB)	20.891	23.656

Table 3. Comparison of Output Images

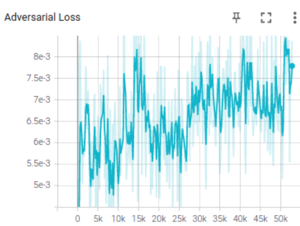
provided in Table 2. The images are shown in Fig2. Comparison of images is shown in Table 3.



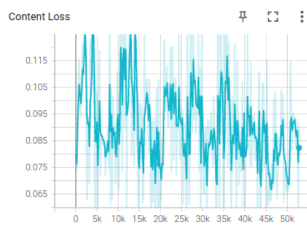
(a) Peceptual loss



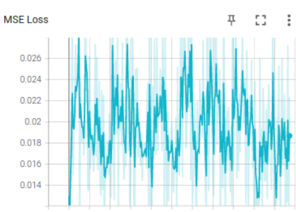
(b) Discriminator loss



(c) Adversarial loss



(c) Content loss



(c) MSE loss

Fig. 1. Training loss plots.

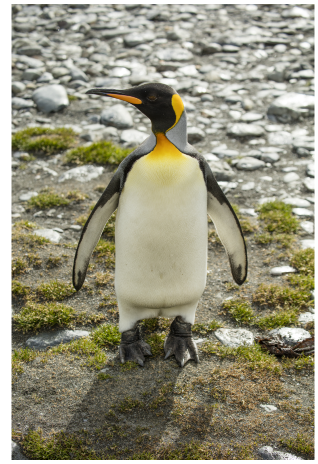
4. CONTRIBUTIONS

The models were implemented from scratch using keras library using tensorflow as backend. I trained SRGAN and proposed a new SRGAN architecture with less parameters which can be used for real time super-resolution. I trained with output size of 256 keeping in view my computational limitations but increasing the size of images while training may improve the results.

5. RESOURCES

- https://data.vision.ee.ethz.ch/cvl/DIV2K/ (DATASET)


(a) LR


(b) HR


(c) Real-Time-SRGAN


(c) SRGAN

Fig. 2. Output Images.

6. REFERENCES

[1] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi, "Photo-realistic single image super-resolution using a generative adversarial network," 2017.

[2] Leonardo Galteri, Lorenzo Seidenari, Marco Bertini, and Alberto Bimbo, *Towards Real-Time Image Enhancement GANs*, pp. 183–195, 08 2019.